

Some notes about matrices, 5

Stephen William Semmes
Rice University
Houston, Texas

As usual, let \mathbf{R} , \mathbf{C} be the real and complex numbers, and for each positive integer n let \mathbf{R}^n , \mathbf{C}^n be the real and complex vector spaces of n -tuples of real and complex numbers, respectively. We write $GL(\mathbf{R}^n)$, $GL(\mathbf{C}^n)$ denote the general linear groups of invertible real and complex-linear transformations on \mathbf{R}^n , \mathbf{C}^n , respectively, and $SL(\mathbf{R}^n)$, $SL(\mathbf{C}^n)$ for the subgroups of $GL(\mathbf{R}^n)$, $GL(\mathbf{C}^n)$ consisting of linear transformations with determinant equal to 1. Also, $O(\mathbf{R}^n)$, $U(\mathbf{C}^n)$ are the subgroups of $GL(\mathbf{R}^n)$, $GL(\mathbf{C}^n)$ of orthogonal and unitary linear transformations on \mathbf{R}^n , \mathbf{C}^n , respectively, which are the linear transformations which preserve the standard Euclidean norm, or, equivalently, the standard inner products on \mathbf{R}^n , \mathbf{C}^n , which is the same as saying that they are invertible linear transformations whose inverses are equal to their adjoints, and $SO(\mathbf{R}^n)$, $SU(\mathbf{C}^n)$ are the special orthogonal and unitary groups, which are the subgroups of $O(\mathbf{R}^n)$, $U(\mathbf{C}^n)$ of linear transformations which also have determinant equal to 1.

Let us write $\mathcal{S}(\mathbf{R}^n)$, $\mathcal{S}(\mathbf{C}^n)$ for the real vector spaces of self-adjoint linear transformations on \mathbf{R}^n , \mathbf{C}^n . We write $\mathcal{S}_+(\mathbf{R}^n)$, $\mathcal{S}_+(\mathbf{C}^n)$ for the open convex cones in $\mathcal{S}(\mathbf{R}^n)$, $\mathcal{S}(\mathbf{C}^n)$ consisting of linear transformations which are positive-definite. Furthermore, we write $\mathcal{M}(\mathbf{R}^n)$, $\mathcal{M}(\mathbf{C}^n)$ for the subsets of $\mathcal{S}_+(\mathbf{R}^n)$, $\mathcal{S}_+(\mathbf{C}^n)$ of linear transformations with determinant equal to 1.

Of course $GL(\mathbf{R}^n)$, $GL(\mathbf{C}^n)$ are open subsets of the vector spaces of all linear transformations on \mathbf{R}^n , \mathbf{C}^n , respectively, which can be defined by the condition that the determinant is nonzero. One can think of $SL(\mathbf{R}^n)$, $SL(\mathbf{C}^n)$ as hypersurfaces in $GL(\mathbf{R}^n)$, $GL(\mathbf{C}^n)$, and indeed they are regular or smooth hypersurfaces because the equation that the determinant be equal to 1 is nondegenerate on these subsets, which is to say that the gradient of the determinant, as a function on the vector space of linear transformations

on \mathbf{R}^n or \mathbf{C}^n , is nonzero at every invertible linear transformation. Similarly, $\mathcal{M}(\mathbf{R}^n)$, $\mathcal{M}(\mathbf{C}^n)$ are smooth hypersurfaces in the open sets $\mathcal{S}_+(\mathbf{R}^n)$, $\mathcal{S}_+(\mathbf{C}^n)$.

There are natural actions of $GL(\mathbf{R}^n)$, $GL(\mathbf{C}^n)$ on $\mathcal{S}(\mathbf{R}^n)$, $\mathcal{S}(\mathbf{C}^n)$, defined by

$$(1) \quad A \mapsto T^* A T$$

when A is a self-adjoint linear transformation on \mathbf{R}^n or \mathbf{C}^n and T is an invertible linear transformation on the same space. If T is an invertible linear mapping on \mathbf{R}^n or \mathbf{C}^n , then this action takes $\mathcal{S}_+(\mathbf{R}^n)$ onto itself or $\mathcal{S}_+(\mathbf{C}^n)$ onto itself, as appropriate. Similarly, if T lies in $SL(\mathbf{R}^n)$ or in $SL(\mathbf{C}^n)$, then this action takes $\mathcal{M}(\mathbf{R}^n)$ onto itself or $\mathcal{M}(\mathbf{C}^n)$ onto itself, as appropriate.

The exponential of a linear transformation A on \mathbf{R}^n or on \mathbf{C}^n is defined by

$$(2) \quad \exp A = \sum_{j=0}^{\infty} \frac{A^j}{j!},$$

and is always an invertible linear transformation on the same space. When A is self-adjoint, $\exp A$ is self-adjoint and positive-definite, and indeed the exponential defines a one-to-one mapping of $\mathcal{S}(\mathbf{R}^n)$, $\mathcal{S}(\mathbf{C}^n)$ onto $\mathcal{S}_+(\mathbf{R}^n)$, $\mathcal{S}_+(\mathbf{C}^n)$, respectively. If we write $\mathcal{S}_0(\mathbf{R}^n)$, $\mathcal{S}_0(\mathbf{C}^n)$ for the linear subspaces of $\mathcal{S}(\mathbf{R}^n)$, $\mathcal{S}(\mathbf{C}^n)$ consisting of self-adjoint linear transformations with trace 0, then the exponential is a one-to-one mapping of $\mathcal{S}_0(\mathbf{R}^n)$, $\mathcal{S}_0(\mathbf{C}^n)$ onto $\mathcal{M}(\mathbf{R}^n)$, $\mathcal{M}(\mathbf{C}^n)$, respectively.

Next we look at some of these objects in terms of vector calculus.

For linear transformations on \mathbf{R}^n or on \mathbf{C}^n , the differential of the exponential mapping at the origin is equal to the identity. Explicitly, for each linear transformation A on \mathbf{R}^n or on \mathbf{C}^n , we have that

$$(3) \quad \left(\frac{d}{dt} \exp(t A) \right)_{t=0} = A.$$

Similarly, for the second differential, if A , B are linear transformations on \mathbf{R}^n or on \mathbf{C}^n , then

$$(4) \quad \left(\frac{\partial^2}{\partial s \partial t} \exp(s A + t B) \right)_{s,t=0} = \frac{1}{2}(A B + B A).$$

Let us define a Riemannian metric on $\mathcal{S}_+(\mathbf{R}^n)$, $\mathcal{S}_+(\mathbf{C}^n)$ as follows. If P is a positive-definite self-adjoint linear transformation on \mathbf{R}^n or on \mathbf{C}^n , then tangent vectors to $\mathcal{S}_+(\mathbf{R}^n)$, $\mathcal{S}_+(\mathbf{C}^n)$ at P are given simply by self-adjoint

linear transformations A, B on \mathbf{R}^n or on \mathbf{C}^n , and we define the inner product of these two tangent vectors at P by

$$(5) \quad \langle A, B \rangle_P = \text{tr}(P^{-1} A P^{-1} B),$$

where $\text{tr } C$ denotes the trace of a linear transformation C . Using standard properties of the trace it is easy to see that this is a positive-definite symmetric bilinear form in A, B , and it depends smoothly on P , so that we get a Riemannian metric on $\mathcal{S}_+(\mathbf{R}^n)$ or on $\mathcal{S}_+(\mathbf{C}^n)$, as appropriate.

Suppose that T is an invertible linear mapping on \mathbf{R}^n or on \mathbf{C}^n , and consider the mapping $\tau = \tau^T$ on $\mathcal{S}_+(\mathbf{R}^n)$ or $\mathcal{S}_+(\mathbf{C}^n)$, as appropriate, defined by

$$(6) \quad \tau(P) = T^* P T.$$

If A, B are self-adjoint linear transformations on \mathbf{R}^n or on \mathbf{C}^n , which we view as tangent vectors to $\mathcal{S}_+(\mathbf{R}^n)$ or $\mathcal{S}_+(\mathbf{C}^n)$ at P , then $d\tau_P(A), d\tau_P(B)$, which are the images of A, B as tangent vectors at P under the differential of the mapping τ , are self-adjoint linear transformations on \mathbf{R}^n or \mathbf{C}^n given by

$$(7) \quad d\tau_P(A) = T^* A T, \quad d\tau_P(B) = T^* B T,$$

and they are viewed as tangent vectors to $\mathcal{S}_+(\mathbf{R}^n)$ or $\mathcal{S}_+(\mathbf{C}^n)$ at $\tau(P)$. One can check that

$$(8) \quad \langle d\tau_P(A), d\tau_P(B) \rangle_{\tau(P)} = \langle A, B \rangle_P,$$

which is to say that the Riemannian metrics on $\mathcal{S}_+(\mathbf{R}^n), \mathcal{S}_+(\mathbf{C}^n)$ are invariant under the actions of $GL(\mathbf{R}^n), GL(\mathbf{C}^n)$ that we have defined.

Of course the usual flat Riemannian metrics are defined as follows. If T is an element of $\mathcal{S}(\mathbf{R}^n)$ or of $\mathcal{S}(\mathbf{C}^n)$, then again two tangent vectors A, B to $\mathcal{S}(\mathbf{R}^n)$ or $\mathcal{S}(\mathbf{C}^n)$ at T are given by self-adjoint linear transformations on \mathbf{R}^n or \mathbf{C}^n , as appropriate, and their inner product in the standard flat Riemannian metric is given by the usual inner product, namely

$$(9) \quad \text{tr } A B.$$

The flatness of this Riemannian metric is reflected in the fact that it does not depend on the point T in the space.

The metrics that we have defined on $\mathcal{S}_+(\mathbf{R}^n), \mathcal{S}_+(\mathbf{C}^n)$ reduce exactly to the standard flat metric at the identity operator I . The differential of the exponential function at the zero operator is equal to the identity mapping,

$$(10) \quad d\exp_0(A) = A,$$

as we have noted previously, and thus the standard flat metric on the tangent space of $\mathcal{S}(\mathbf{R}^n)$ or $\mathcal{S}(\mathbf{C}^n)$ at 0 agrees with the metric that we have defined on $\mathcal{S}_+(\mathbf{R}^n)$ or $\mathcal{S}_+(\mathbf{C}^n)$, respectively, at the identity operator I with respect to the correspondence given by the differential of the exponential function. We would like to show that in fact this agreement works to another term in the Taylor expansion.

Basically this means that if T is a self-adjoint linear transformation on \mathbf{R}^n or on \mathbf{C}^n , which we think of as being near the origin, and if A, B are two self-adjoint linear transformations on the same space, which we think of as tangent vectors to $\mathcal{S}(\mathbf{R}^n)$ or $\mathcal{S}(\mathbf{C}^n)$ at T , then the images of A, B under the differential of the exponential map at T have inner product with respect to the Riemannian metric defined above at the exponential of T which agrees to second order with the inner product of A in the standard flat metric on $\mathcal{S}(\mathbf{R}^n)$ or $\mathcal{S}(\mathbf{C}^n)$. Explicitly, this means that

$$\begin{aligned} (11) \quad & \langle d \exp_T(A), d \exp_T(B) \rangle_{\exp T} \\ (12) \quad & = \text{tr}(\exp(-T))(d \exp_T(A))(\exp(-T))(d \exp_T(B)) \end{aligned}$$

agrees with

$$(13) \quad \text{tr } A B$$

up to terms of order $O(\|T\|^2)$. This is not difficult to check, using the facts that $\exp(-T) = I - T + O(\|T\|^2)$,

$$(14) \quad d \exp_T(A) = A + \frac{1}{2}(T A + A T) + O(\|T\|^2),$$

and similarly for B .

As a consequence, if C is any self-adjoint linear transformation on \mathbf{R}^n or \mathbf{C}^n , then

$$(15) \quad \exp(t C)$$

satisfies the equation for geodesics in the space of positive-definite linear transformations at $t = 0$. Basically this is because the line $t C$ satisfies the equation for geodesics in the flat space of self-adjoint linear transformations at the origin. Of course we are also using the fact that the standard flat metric on the space of self-adjoint linear transformations around 0 agrees with the Riemannian metric on the space of positive-definite metrics around the identity operator I with respect to the exponential mapping as well as they do.

In fact, we get that

$$(16) \quad \exp(tC)$$

satisfies the equation for geodesics in the space of positive-definite matrices for all real numbers t . For a fixed real number t_0 , the statement that this curve satisfies the equation for geodesics at t_0 is equivalent to the statement that $\exp((t_0 + t)C)$ satisfies the equation for geodesics at $t = 0$. This is in turn equivalent to the statement that $T^* \exp(tC) T$ satisfies the equation for geodesics at $t = 0$ when $T = \exp(t_0 C/2)$, and this follows from the fact that $\exp(tC)$ satisfies the equation for geodesics at $t = 0$, since transformations of the form $P \mapsto T^* P T$ are isometries on the spaces of positive-definite matrices and therefore preserve geodesics.

Thus exponentials of straight lines through the origin in the spaces of self-adjoint linear transformations give rise to geodesics through the identity operator I in the corresponding spaces of positive-definite linear transformations. By standard results about uniqueness of initial value problems for ordinary differential equations this accounts for all of the geodesics in the spaces of positive definite linear transformations which pass through the identity operator I . Using mappings of the form $P \mapsto T^* P T$ one can move these curves to other places and thereby account for all of the geodesics in the spaces of positive-definite linear transformations.